

CONNECTING THE CONTINUUM

*Longitudinal Data Systems in
North Carolina*

June 2019

Javaid Siddiqi, Ph.D., President & CEO

Patrick C. Sims, Director of Policy and Research

Allison L. Goff, Policy Analyst

ACKNOWLEDGMENTS

We extend our deep appreciation to the partners of the North Carolina Education Longitudinal Data System for sharing their time, insights, and detailed feedback throughout the writing of this report. Thanks to the North Carolina Community College System, North Carolina Department of Public Instruction, North Carolina Government Data Analytics Center, North Carolina Department of Health and Human Services, North Carolina Independent Colleges and Universities, the University of North Carolina System, members and staff of the North Carolina General Assembly, and the Office of Governor Roy Cooper. Additionally, we extend our gratitude to the Bill & Melinda Gates Foundation for their support of this report.

We would also like to thank the teams at Washington's Education Research Data Center and the Kentucky Center for Statistics for sharing their perspective on longitudinal data systems, and lessons learned from their own statewide longitudinal data systems.

Additional thanks for the collaborative support of the team at The Hunt Institute including Michele Jordan, Jon-Taidje Joyner, Jennifer Landes, James Mikolowsky, and to Hannah Bartlebaugh and Allie Jaarsma for their research and editing support.

ABOUT THE HUNT INSTITUTE

An affiliate of the Duke University Sanford School of Public Policy, The Hunt Institute is a recognized leader in the movement to transform public education. Marshaling expertise from a nationwide partner network since it was established in 2001, The Institute brings together people and resources that help build and nurture visionary leadership and mobilize strategic action for greater educational outcomes and student success. For more information, please visit: <http://www.hunt-institute.org/>

FOREWORD

Policymakers today are uniquely positioned to leverage data to inform decision-making in unprecedented ways. More data are being collected than ever before and analytical methods continue to advance. North Carolina, in particular, has a rich history of valuing and prioritizing the collection of relevant education data and information, giving it a head start on creating a modern statewide longitudinal data system (SLDS).

But what does a high-quality SLDS look like and how can it support our complex education system?

We know that learning, and other key factors that influence it such as health and nutrition, begins even before birth. And we also know that in today's rapidly evolving economy, learning must be a continuous endeavor that spans an individual's career. As a result, it is critical that an SLDS incorporates a broad array of systems, including early childcare, pre-K, K-12, community colleges, private and public four-year postsecondary institutions, and state agencies who own workforce data.

Through conversations with SLDS leaders across the country, we have seen that a lack of collaboration across state agencies, including limited data sharing, poor coordination of data consistency, and lack of investment in agency systems can severely limit the success of an SLDS.

However, policymakers are in the unique position to coalesce these stakeholders to build a system that can be sustained across multiple administrations.

Additionally, legislators need to ensure that system governance is well structured, building trust and transparency among and between partners. The governing body serves a pivotal role in ensuring the SLDS is functional. Building a governance structure that is transparent, codified in legislation, and fully funded will ensure that the system functions efficiently and is flexible enough to grow over time to meet the evolving needs of the state.

As we move forward, we need to think critically about how to use the data system in a way that informs policy decisions. How, for example, can district leaders use this system to inform their work? How can institutions use this information to improve educator preparation programs?

There are countless ways that a high-quality system can be leveraged to improve outcomes for all young people. We look forward to helping drive progress on this important work.



Javaid Siddiqi, Ph.D.
President & CEO

GUIDING QUESTIONS FOR POLICYMAKERS

This report explores the history, purpose, and future of statewide longitudinal data systems. While the focus of this report is on North Carolina, it also draws on national lessons from other states, especially Kentucky and Washington. North Carolina has already made significant progress on building out its data systems but will require additional work in order to meet the education and research needs of the state.

As policymakers take a closer look at statewide longitudinal data systems (SLDS), there are a number of questions they should consider when determining the effectiveness and functionality of the Systems. The following questions should help guide policymakers as they consider ways in which they can support the creation and use of education data systems.

1 What questions do we need to answer in order to make progress toward our state's education goals?

Many of the questions that can help inform and shape policy decisions can only be answered by combining data from multiple sources and over longer periods of time.

2 Which audiences will benefit from data provided through the System?

A system should serve policymakers, state-level agencies, system partners, researchers, and families.

3 What partners are engaged in data sharing?

Policymakers should consider whether all educational entities are sharing data with the SLDS, including early intervention, early childcare, pre-K, K-12, community colleges, private and public four-year postsecondary institutions, and state agencies who own workforce data.

4 Does the SLDS have a robust governance structure in place to provide necessary oversight and coordination?

The System needs a clear governing body, codified in legislation, which has worked closely with partners in order to build and operate a system that produces timely and accurate data, while ensuring that student privacy is protected.

5 What are the staffing and technological requirements needed to reliably collect, maintain, and submit data? What personnel are needed to ensure effective governance and communication?

In order to function effectively, each partner entity must have sufficient human capital to submit, clean, and analyze data. Additionally, the governing body must have analysts to complete quality control checks, process requests, and coordinate partners.

6 How is the System funded, and is that funding sustainable?

The System must be funded through recurring funds from the legislature. Additionally, each partner agency must have sufficient recurring funds to support infrastructure and human capital needs.

7 How will stakeholders access the information in the System?

The System should produce sets of de-identified data that can be analyzed by researchers, regular aggregate reports to serve policymakers, and a public-facing interface with data presented in a way that is easily understood.

8 What steps are being taken to ensure that data privacy is protected?

Efforts to ensure data privacy must be at the forefront of all statewide data efforts.

CONNECTING THE CONTINUUM

Longitudinal Data Systems in North Carolina

01 INTRODUCTION

01 STATEWIDE LONGITUDINAL DATA SYSTEMS

01 History & Funding

03 Structure and Organization

05 Key Characteristics of a High-Quality SLDS

06 LANDSCAPE ANALYSIS: NORTH CAROLINA

06 A Strong History of Data

08 The Creation of the NCSchoolWorks Longitudinal Data System

10 North Carolina Education Longitudinal Data System

*11 Challenges Facing the North Carolina Education Longitudinal
Data System*

11 Future of Longitudinal Data Systems in North Carolina

12 STATE SLDS CASE STUDIES: KENTUCKY & WASHINGTON

12 Kentucky

12 History and Formation

12 System Overview

13 Governance

13 Use and Impact

14 Next Steps

14 Keys to Success

14 Challenges

15 Washington

15 History and Formation

15 System Overview

15 Governance

16 Use and Impact

16 Next Steps

17 Keys to Success

17 Challenges

18 GUIDING QUESTIONS FOR POLICYMAKERS

20 GLOSSARY

21 CITATIONS

Introduction

The rise of data collection in the public and private sectors has presented policymakers with an unprecedented opportunity to leverage vast amounts of data to create more efficient systems, improve program performance, and drive equitable outcomes. Empowering policymakers and practitioners with functional access to accurate data has the potential to change the way we deliver education. States across the country are currently seeking ways to build systems that effectively collect, compile, and analyze these data to answer complex questions.

With a strong history of developing and utilizing data systems that span the fields of education, health, and the workforce, North Carolina could lead the nation in the development and use of longitudinal data systems. Through the strong commitment and relentless efforts of the North Carolina Department of Commerce, North Carolina Community College System (NCCCS), North Carolina Department of Health and Human Services (NCDHHS), North Carolina Independent Colleges and Universities (NCICU), North Carolina Department of Public Instruction (NCDPI), and the University of North Carolina System (UNC), significant progress has been made in creating longitudinal data systems in North Carolina. However, North Carolina continues to struggle to connect disparate data sources to one another and ensure consistency across systems, keeping the state from fully leveraging data to regularly inform policy and practice in a meaningful way.

Most state agencies collect data independently of one another and typically do not have structures in place to guarantee alignment with the types of data points collected and how each are defined. Though each agency has some capacity to use its own data to analyze performance, larger policy questions can be answered when the data are compiled across agencies and time. These can be questions of practice that guide the work of principals and teachers, such as “What percentage of my secondary students required remedial coursework in my subject area at the postsecondary level?” They may also be questions of policy that impact broader programmatic and budget decisions, such as “How likely are students who are reading on grade level by third grade to complete a postsecondary degree or credential?” To answer these questions, an analyst needs student-level data from multiple government agencies over a number of years—something that has proven to be challenging in most states.

As data collection has grown in prominence, privacy concerns have also risen. Along with the benefits that come with this data collection—such as the personalization of services and greater efficiencies—many constituents are concerned about privacy, and policymakers need to ensure any data systems are built with security in mind.

Statewide Longitudinal Data Systems

History & Funding

The first major push for statewide longitudinal data systems (SLDS) came through the *Educational Technology and Assistance Act of 2002*, which created the first competitive federal grant process to support state education agencies in the development of these systems.¹ These grants were first distributed by the Institute of Education Sciences (IES) in 2006, with subsequent grants being awarded in 2007, 2009, 2012, and 2015. The largest block of grants was awarded as a part of the *American Recovery & Reinvestment Act of 2009* (ARRA), which supplemented the 2009 grants with an additional \$230 million. In total, over \$700 million has been awarded to 47 states, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands through the SLDS grants.

Additional grant funding has been awarded through the Race to the Top competitive grant process,

which encouraged states to make policy reforms that would enhance standards and assessments, improve the collection and use of data, increase teacher effectiveness, and turn around struggling schools. More specifically, ARRA required that states wishing to receive Race to the Top funds must have made significant progress in establishing longitudinal data systems, making this another mechanism to incentivize states to improve their data systems.

Figure 1 | SLDS Federal Grant Funding by Year²

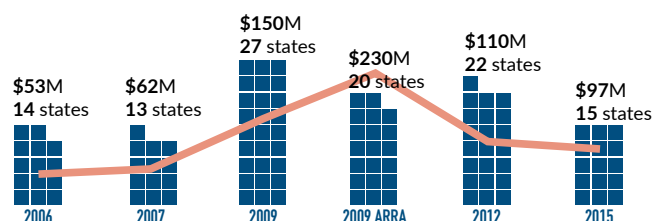


Figure 2 | SLDS Federal Grant Program Funding by State³

STATE	2006	2007	2009	2009 ARRA	2012	2015	TOTAL	
Texas	\$0	\$0	\$7,879,785	\$18,195,078	\$0	\$6,972,522	\$33,047,385	
Pennsylvania	\$4,008,875	\$0	\$6,103,000	\$14,284,020	\$0	\$6,999,928	\$31,395,823	
Washington	\$0	\$0	\$5,941,887	\$17,341,871	\$0	\$6,992,452	\$30,276,210	
Illinois	\$0	\$0	\$8,999,956	\$11,869,819	\$0	\$7,000,000	\$27,869,775	
Wisconsin	\$3,081,000	\$0	\$5,552,270	\$13,809,040	\$0	\$5,242,866	\$27,685,176	
New York	\$0	\$0	\$7,844,313	\$0	\$19,670,975	\$0	\$27,515,288	
Massachusetts	\$0	\$0	\$5,993,464	\$12,972,730	\$0	\$6,999,761	\$25,965,955	
Virginia	\$0	\$6,054,394	\$0	\$17,537,564	\$0	\$0	\$23,591,958	
Minnesota	\$3,272,448	\$0	\$0	\$12,411,777	\$0	\$6,992,025	\$22,676,250	
Maryland	\$5,690,718	\$0	\$5,990,186	\$0	\$3,963,473	\$6,990,361	\$22,634,738	
Colorado	\$0	\$4,244,519	\$0	\$17,409,117	\$0	\$0	\$21,653,636	
South Carolina	\$5,795,603	\$0	\$0	\$14,890,261	\$0	\$0	\$20,685,864	
Utah	\$0	\$4,561,763	\$0	\$9,617,736	\$0	\$6,497,783	\$20,677,282	
Michigan	\$3,000,000	\$0	\$5,517,228	\$10,624,964	\$0	\$0	\$19,142,192	
Kentucky	\$5,780,275	\$0	\$2,878,373	\$0	\$3,633,928	\$6,634,741	\$18,927,317	
Oregon	\$0	\$4,705,977	\$3,696,615	\$10,475,997	\$0	\$0	\$18,878,589	
Arkansas	\$3,328,503	\$0	\$4,967,991	\$9,832,689	\$0	\$0	\$18,129,183	
Mississippi	\$0	\$0	\$3,387,308	\$7,569,716	\$0	\$6,588,210	\$17,545,234	
North Dakota	\$0	\$0	\$6,723,090	\$0	\$3,943,898	\$6,475,690	\$17,142,678	
Kansas	\$0	\$3,834,796	\$3,911,792	\$9,060,442	\$0	\$0	\$16,807,030	
Florida	\$1,577,602	\$0	\$2,450,000	\$9,975,288	\$0	\$0	\$14,002,890	
Ohio	\$5,670,100	\$0	\$2,945,000	\$5,135,883	\$0	\$0	\$13,750,983	
Hawai'i	\$0	\$0	\$3,477,053	\$0	\$3,386,693	\$6,642,010	\$13,505,756	
Montana	\$0	\$0	\$5,798,457	\$0	\$3,977,861	\$3,483,163	\$13,259,481	
Iowa	\$0	\$0	\$8,777,459	\$0	\$3,747,281	\$0	\$12,524,740	
Arizona	\$0	\$5,954,518	\$0	\$0	\$4,966,706	\$0	\$10,921,224	
Maine	\$0	\$3,227,231	\$0	\$7,315,000	\$0	\$0	\$10,542,231	
Tennessee	\$3,226,313	\$0	\$0	\$0	\$0	\$6,917,059	\$10,143,372	
Nevada	\$0	\$5,999,975	\$0	\$0	\$3,999,990	\$0	\$9,999,965	
D.C.	\$0	\$5,738,500	\$0	\$0	\$4,000,000	\$0	\$9,738,500	
North Carolina	\$0	\$6,000,000	\$0	\$0	\$3,639,543	\$0	\$9,639,543	
California	\$3,255,445	\$0	\$6,000,000	\$0	\$0	\$0	\$9,255,445	
Indiana	\$0	\$5,188,260	\$0	\$0	\$3,965,160	\$0	\$9,153,420	
Idaho	\$0	\$0	\$5,916,520	\$0	\$3,101,632	\$0	\$9,018,152	
Missouri	\$0	\$0	\$8,967,686	\$0	\$0	\$0	\$8,967,686	
Georgia	\$0	\$0	\$8,942,640	\$0	\$0	\$0	\$8,942,640	
Rhode Island	\$0	\$0	\$4,667,933	\$0	\$4,000,000	\$0	\$8,667,933	
New Hampshire	\$0	\$3,176,272	\$0	\$0	\$4,989,391	\$0	\$8,165,663	
Nebraska	\$0	\$3,468,335	\$0	\$0	\$4,361,534	\$0	\$7,829,869	
Alaska	\$3,506,757	\$0	\$0	\$0	\$4,000,000	\$0	\$7,506,757	
Oklahoma	\$0	\$0	\$0	\$0	\$4,997,082	\$0	\$4,997,082	
Vermont	\$0	\$0	\$0	\$0	\$4,947,261	\$0	\$4,947,261	
West Virginia	\$0	\$0	\$0	\$0	\$4,798,697	\$0	\$4,798,697	
Delaware	\$0	\$0	\$0	\$0	\$4,616,250	\$0	\$4,616,250	
Connecticut	\$1,500,714	\$0	\$2,937,416	\$0	\$0	\$0	\$4,438,130	
Louisiana	\$0	\$0	\$4,056,510	\$0	\$0	\$0	\$4,056,510	
New Jersey	\$0	\$0	\$0	\$0	\$3,989,175	\$0	\$3,989,175	
South Dakota	\$0	\$0	\$0	\$0	\$3,033,792	\$0	\$3,033,792	
Alabama	\$0	\$0	\$0	\$0	\$0	\$0	\$0	
New Mexico	\$0	\$0	\$0	\$0	\$0	\$0	\$0	
Wyoming	\$0	\$0	\$0	\$0	\$0	\$0	\$0	
TOTAL	\$52,694,353	\$62,154,540	\$150,323,932	\$230,328,992	\$109,730,322	\$97,428,571	\$702,660,710	

Structure and Organization

Statewide longitudinal data systems connect individual-level data over time and across agencies. While most states refer to these systems as SLDSs, in North Carolina the System is known as the Education Longitudinal Data System (ELDS). These data systems allow researchers and policymakers to follow a student's academic career, draw inferences about what types of programs and services are working well, and identify disparities in performance among subgroups. Education-focused SLDSs are often referred to as P-20W systems, which cover pre-Kindergarten (P), K-12 and postsecondary (20), and workforce (W) data.

All 50 states and the District of Columbia are able to connect data between systems in some capacity, and 38 states and the District of Columbia are able to connect at least two of the four core systems (early learning, K-12, postsecondary, and workforce). Only 17 states and the District of Columbia have developed a full P-20W system that connects all four core systems.⁴

There are two main types of SLDSs—centralized and federated—though, in practice, there are many variations that fall somewhere in between the two approaches. North Carolina's system, for example, will pull data from both federated and centralized systems. Each type has the capacity to collect the same data and facilitate the same types of analyses, but they differ in how the data are compiled and where the databases are housed. According to recent analysis, only 11 of the 38 systems are federated.⁵

Data are typically matched across systems based on a set of student characteristics, including name, date of birth, gender, other demographic data, and, in some instances, social security numbers.⁶ Many states also assign a unique identifier, or UID, to each student, which follows them throughout their academic career. The use of a UID ensures that student data are not associated with personally identifiable information once in the system. This UID is consistent across systems and would, for example, follow a student as they move from early childhood services to K-12, community college, and university. UIDs can be especially helpful in a federated system, as they can facilitate quick and accurate matching across systems. The consistent and accurate assignment of a UID is critical to the success of systems that use them.

Figure 3 | Centralized vs Federated Data Systems

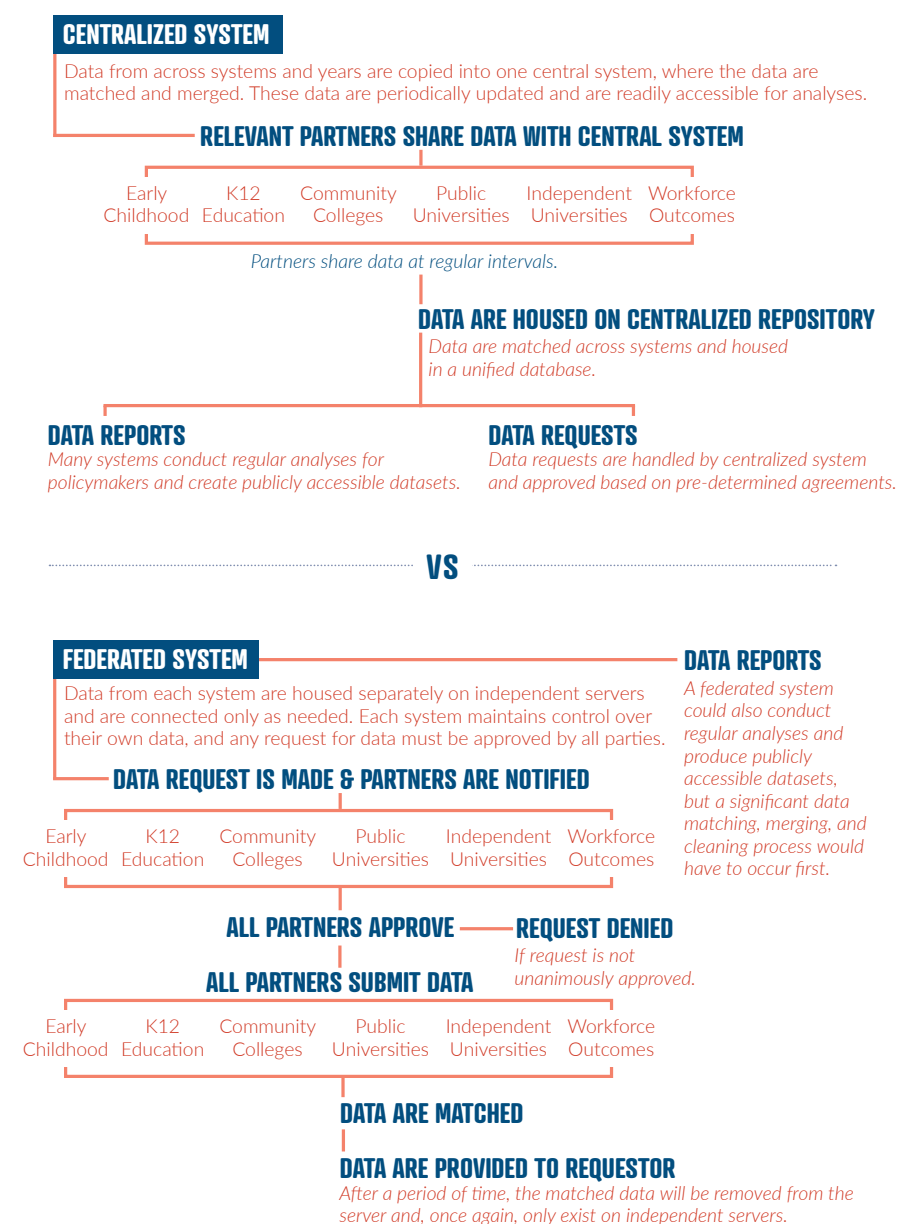


Figure 4 | Data That Can Be Collected in a North Carolina P-20W System

Most states and state agencies are already collecting a wealth of information on students, including the more detailed data points listed below, many of which are specific to the North Carolina system. A robust P-20W system would cover a broad range of student, teacher, and school characteristics, as well as data on student performance from agencies—starting with early childhood programs and services going through workforce outcomes. Consistent and uniform reporting on each of the variables will allow for reliably connecting these data for meaningful analysis.

Early Childhood	<ul style="list-style-type: none">› State pre-Kindergarten› State-funded Head Start› Federally-funded Head Start› Subsidized child care› Developmental screening data› Vital records data› Health screening, immunization, and well-child visit data		<ul style="list-style-type: none">› Enrollment in private pre-Kindergarten, ECE or daycare centers› Early intervention› Early childhood workforce› Pre-Kindergarten special education› Home visiting	<ul style="list-style-type: none">› Race/Ethnicity› Gender› Special education/ Gifted status› FRL status/Other SES indicator› Geographic location (urban/ rural/suburban)› Teacher-student data link› Student UID› Highest level of parental educational attainment	
Kindergarten	<ul style="list-style-type: none">› Kindergarten entry assessment› mCLASS data		<ul style="list-style-type: none">› Absenteeism› Behavioral (short-term suspension, long-term suspension, expulsion)› Grade retention› Information on untested students› Teacher data› English learner status› IEP or 504› School transfer		
Grade 1	<ul style="list-style-type: none">› mCLASS data				
Grade 2	<ul style="list-style-type: none">› mCLASS data				
Grade 3	<ul style="list-style-type: none">› EOG ELA/Reading› EOG Math› EOG Science				
Grade 4					
Grade 5					
Grade 6					
Grade 7	<ul style="list-style-type: none">› EOG ELA/Reading› EOG Math› EOG Science				
Grade 8					
Grade 9					<ul style="list-style-type: none">› AP score
Grade 10					<ul style="list-style-type: none">› AP score› PreACT score
Grade 11	<ul style="list-style-type: none">› ACT score› AP/IB score› Dual enrollment	<ul style="list-style-type: none">› Remedial coursework› Graduation status› Dropout status› CTE enrollment› Course enrollment and grades› EOC English II› EOC NC Math 1› EOC NC Math 3› EOC Biology			
Grade 12	<ul style="list-style-type: none">› ACT score› AP score› CTE enrollment› Dual enrollment				
Community College & University	<ul style="list-style-type: none">› Institutions attended› Financial aid received› Course enrollment and grades› Degree or certificate earned› Remedial coursework				<ul style="list-style-type: none">› Graduation› Dropout› Retention› Persistence
Workforce	<ul style="list-style-type: none">› Employment status› Wages› Unemployment benefits data				

Figure 5 | America COMPETES Act: 12 Essential Elements of an SLDS by Academic Level

In 2007, the America COMPETES Act called for alignment across the education continuum, including improvement of longitudinal data systems. The Act identified the following 12 essential elements of an SLDS.⁷ These data points range from student achievement and demographics to remedial coursework and the ability to match students with teachers.

Pre-K Through 12	Postsecondary	Pre-K Through Postsecondary
<ul style="list-style-type: none"> › Student transition success data and enrollment in remedial education › Data to assess adequate preparation for postsecondary success 	<ul style="list-style-type: none"> › Unique student identifier that maintains confidentiality › Student-level enrollment, demographic, and participation › Student-level data on completion, transfer, and dropout › Ability of system to communicate with other systems › Audit system to assess data quality 	<ul style="list-style-type: none"> › Annual test records › Data on untested students by grade and subject › Teacher identifier that can match teachers with students › Student-level transcript information › Student-level college readiness test scores

Key Characteristics of a High-Quality SLDS

Statewide longitudinal data systems must possess a number of characteristics and require a favorable policymaking environment in order to provide high-quality, useful education data to a wide variety of stakeholders. Among other characteristics, the following outlines some of the most critical elements to a successful SLDS.

- ▶ **Widespread Policymaker Support:** In order to effectively drive change, advocates for improving a state's education data system must build the momentum needed to develop support among elected state officials and state agency staff. State legislatures have the capacity to codify SLDS policy in ways that can sustain a system across multiple administrations, and the governor can serve as a coordinating force. Policymakers are in the unique position to have the leverage necessary to compel collaboration.
- ▶ **Meaningful Agency Collaboration:** Many SLDSs have struggled due to lack of collaboration across state agencies. This comes in many forms, including limited data sharing, poor coordination of data consistency, and lack of investment in agency systems. It is critical that all agencies see the potential benefits of the System and that they are committed to transparency and building trust.
- ▶ **A High-Quality Governance Structure:** Governance structures are responsible for making critical decisions about how student data are collected, linked, reported, and protected. In addition, SLDSs require a significant amount of maintenance and ongoing updates. To keep the System functioning, it is critical that structures are set up to ensure regular communication.
- ▶ **Robust Privacy and Security Protocols:** Concerns about privacy are shared by a variety of stakeholders, including parents, schools, agencies, and policymakers. These concerns are justifiable but can be properly addressed using the appropriate protections in data management, such as de-identification and data minimization.
- ▶ **Standardization of Data:** With multiple agencies gathering data, it is critical that there is a common understanding of how each variable should be defined. Small differences in definitions can make analyses unreliable and tasks burdensome for researchers. Agencies need to build their capacity to ensure the reliability of their own data and coordinate with other agencies to ensure accurate data matching.
- ▶ **Adequate Technical Infrastructure & Dedicated Human Capital:** Each SLDS partner must have sufficient staff and technology to allow for efficient and accurate data recording and transferring. Many states have failed to supplement federal start-up money with recurring state funds in order to maintain systems, staff, and technological capacities.
- ▶ **Sustainable Internal Leadership:** The institutional knowledge possessed by staff is valuable, and the loss of that knowledge due to staff transitions can damage the System as a whole. Too often, agencies rely on one individual to lead the work, build the requisite knowledge, and develop the necessary relationships. Agencies need to be thoughtful about structuring this work so that it is not a challenge to sustain it in the wake of staff turnover or after electoral transitions.

Landscape Analysis: North Carolina

A Strong History of Data

North Carolina's education system has long served as a resource for educational research due in large part to the presence of data systems capturing student, teacher, and school performance data going back to the mid-1990s. North Carolina first created a longitudinal data system with the launch of the Common Follow-up System (CFS) in 1992. Originally created by state agencies to capture outcomes of public education, employment, and training programs, the CFS was written into statute by the North Carolina General Assembly in 1995.⁸ The CFS currently includes data from the North Carolina Department of Commerce, North Carolina Department of Health and Human Services (NCDHHS), North Carolina Department of Public Instruction (NCDPI), North Carolina Department of Public Safety, North Carolina Community College System (NCCCS), and the 17-campus University of North Carolina System (UNC).⁹

In 2000, the North Carolina Education Research Data Center was established through a partnership between NCDPI and the Duke University Center for Child and Family Policy.¹⁰ Researchers from nonprofit and university settings, as well as government agencies, can pay to access more than 20 years of data on North Carolina's public schools, students, and teachers. In addition to this private source of educational data, the state has a long history of publicly available workforce and postsecondary data as well. This commitment to gathering data laid the groundwork for what would eventually become North Carolina's Education Longitudinal Data System (ELDS).

In an effort to create the infrastructure needed to analyze different types of data over time and improve the quality and accessibility of pre-K through secondary educational data, NCDPI applied for and received a \$6 million grant from the U.S. Department of Education as part of the Statewide Longitudinal Data Systems Grant Program in 2007. Under this grant, NCDPI developed the Common Education Data Analysis and Reporting

Figure 6 | The North Carolina Education Longitudinal Data System

NORTH CAROLINA EDUCATION LONGITUDINAL DATA SYSTEM (ELDS)			
	Early Childhood Integrated Data System (ECIDS)	NCSchoolWorks	Common Follow-up System (CFS)
NC Department of Health and Human Services (NCDHHS)	✓		✓
NC Head Start/Early Head Start	✓		
NC Department of Public Instruction (NCDPI)	✓	✓	✓
NC Independent Colleges and Universities (NCICU)*		✓	
NC Community College System (NCCCS)		✓	✓
UNC System		✓	✓
NC Commerce		✓	✓

**Data from North Carolina's independent colleges and universities are not currently available for requests pending the creation of formal rules and memoranda of understanding between each independent college or university and the partner organizations.*

System (CEDARS), which included the creation of a Unique Statewide Identifier (UID) for both students and teachers. Use of a UID allowed for staff- and student-level data to be matched across the K-12 continuum.

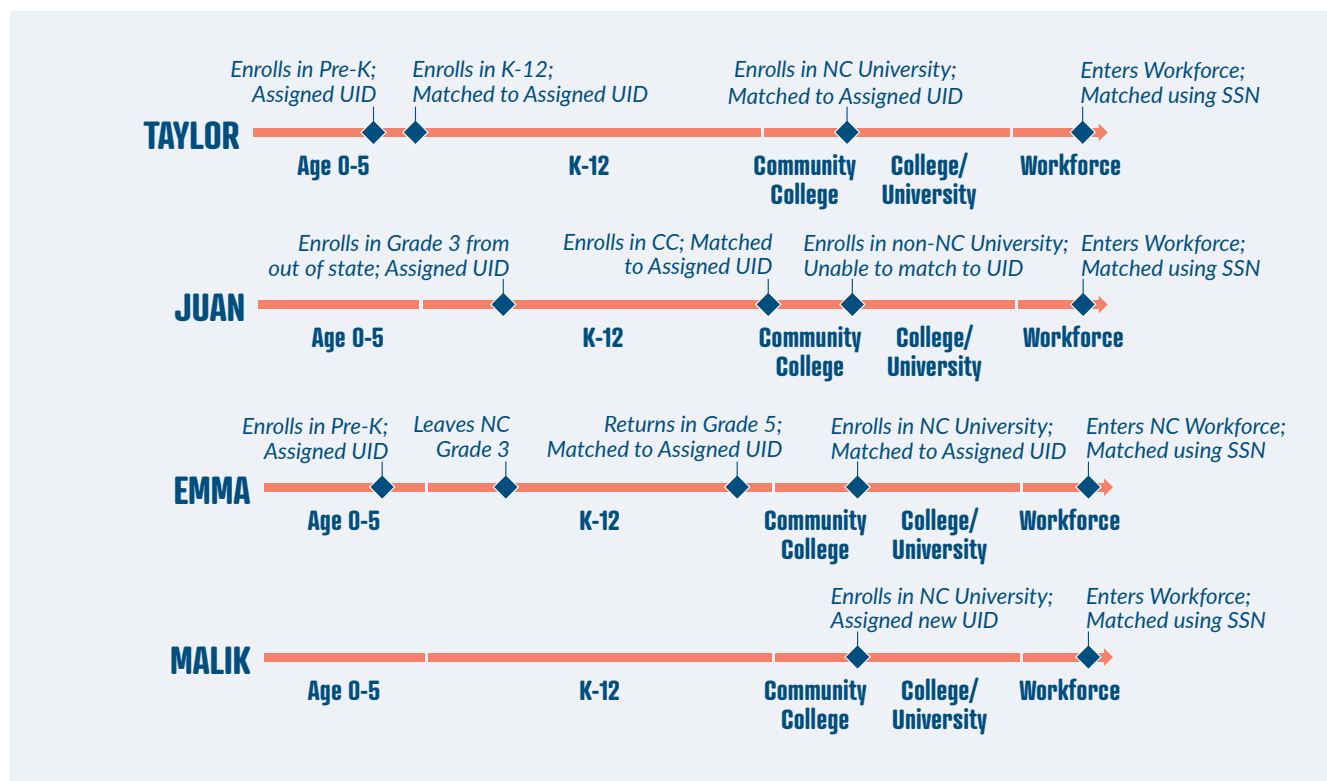
The effectiveness of CEDARS was supported by a 2010 statewide mandate from NCDPI requiring all North Carolina public schools to utilize the PowerSchool technology platform as an operational data store. Of critical importance, PowerSchool utilizes the UID interface for students and teachers, improving data validity for the K-12 sector.

A student enrolled in a North Carolina K-12 public school is assigned a UID upon enrollment. This is a mandatory step in the school enrollment process across K-12 schools in the state. Ensuring that a student's UID successfully travels with them to a postsecondary setting has proven to be a challenge. Part of the difficulty lies in the timing of when an institution checks to determine if a student already has a UID. Students transferring between K-12 schools are generally matched or assigned new UIDs without issue because all North Carolina schools match UIDs locally (i.e., in person) and use the same technology platform for enrollment across schools.

When a student enrolls in a school for the first time, the System automatically checks to determine if that student has a UID and either matches the pre-existing UID to that student or notifies the individual completing enrollment that there is a near match (i.e., some of the demographic details indicate a match, but others do not). In cases of a near match, the use of local UID matching allows the near match to be resolved immediately, in person, by clarifying any inconsistencies in the student's profile. In instances where the student has never been enrolled, a new UID is assigned in real-time.

In 2011, North Carolina was awarded a \$6.9 million Race to the Top – Early Learning Challenge Grant that included the creation of the North Carolina Early Childhood Integrated Data System (NCECIDS). This birth-to-five system is administered by the NCDHHS and includes data on early childhood education, health, and social services.¹¹ Importantly, NCECIDS utilizes the UID software, making it possible to integrate early childhood data into the K-12 longitudinal data system. The North Carolina General Assembly has approved recurring state funding to support the operation and maintenance of NCECIDS.

Figure 7 | Examples of Unique Identifier (UID) Assignment and Use



NCDHHS is currently working to increase the data sources and data points available through NCECIDS. Data from the Temporary Assistance for Needy Families (TANF) program are anticipated to be fully integrated into NCECIDS by the end of 2019. A pilot program to integrate data regarding children enrolled in Head Start is currently in progress for a subset of grantees in preparation for all Head Start data to be included. Finally, NCDHHS is working to standardize data collection from home visiting programs in preparation for those data to be integrated into NCECIDS in the future.

The Creation of the NCSchoolWorks Longitudinal Data System

With the creation of early childhood and K-12 longitudinal data systems, and the continued use of the CFS capturing postsecondary and workforce data, North Carolina was primed to create a P-20W system that would span each pre-existing system. NCDPI led this effort and was awarded a \$3.6 million grant from the U.S. Department of Education to develop a federated, P-20W statewide longitudinal data system. The system would later be named NCSchoolWorks. The P-20W SLDS collaboration includes NCDPI, UNC, NCCCS, NCICU, and the North Carolina Department of Commerce Division of Employment Security.

The efforts of the partner entities were signed into law in 2012 with the passage of *House Bill 964*, creating the North Carolina Longitudinal Data System. The enacted bill became North Carolina General Statute Chapter 116E. Per the bill, the System “is a statewide data system that contains individual-level student and workforce data from all levels of education and the State’s workforce” located administratively within, but independently of, NCDPI.¹² This statute has been used to provide guidance to both NCSchoolWorks and the ELDS, which will be discussed later in this section.

The system was originally governed by an 18-member North Carolina Longitudinal Data System Board which included the agency leads from the partners, as well as representatives appointed by the legislature. The state statute was amended in 2016 to transfer responsibility over the System from the Board to North Carolina’s Governmental Data Analytics Center (GDAC).¹³ Those responsibilities include development

of an implementation plan, provision of general oversight and direction, approval of an annual budget, and assurance that privacy and data security are in line with other laws. Additionally, an advisory committee is to advise the GDAC on data quality and data validity. While membership is not included in the statute, the advisory committee consists of representatives of each of the partner agencies, as well as GDAC, and the Office of the Governor.

The statute goes on to outline the duties and functions of the System, which include serving as a data broker, compliance with privacy laws and policies, and facilitation of data requests for state and federal education reporting and public information requests. The statute mandates that data accessed through the System must only be accessed by authorized staff, must be de-identified, and can only be used in aggregate form when used in reports or when shared as part of a data request.

By the end of the grant period in 2017, the System was found to have met the objectives outlined in the grant.

The System is currently being tested, and while it is reported to work, it is only accessible to the partners. At present, the system is able to produce data from NCDPI, the UNC System, and the Department of Commerce. It is possible to match data from NCCCS that is pulled from the CFS and there continues to be efforts to improve infrastructure at NCCCS such that those data can be pulled directly from the NCSchoolWorks. Early childhood data from NCECIDS cannot be joined to the NCSchoolWorks at this time. Additionally, data from North Carolina’s independent colleges and universities are not currently available for requests pending the creation of formal rules and memoranda of understanding between each independent college or university and the partner organizations.

In 2018, North Carolina Governor Roy Cooper reconvened the North Carolina Education Cabinet, which includes the State Superintendent of Public Instruction, the President of NCCCS, the President of the UNC System, the President of NCICU, the Secretary of NCDHHS, and the Secretary of Commerce. The Governor charged the Cabinet with convening a working group from the NCSchoolWorks partners and GDAC to continue the development of the broader ELDS with GDAC.

Figure 8 | Proposed NCSchoolWorks Data Request Process

RESEARCHER MAKES REQUEST

- » Submits information about research request including research question, IRB approval, and intended use of data.
- » Identifies which data are requested by selecting relevant Common Education Data Standards (CEDS).
- » Identifies partner entity from which CEDS are requested.

RELEVANT PARTNERS ARE NOTIFIED

NCDPI NCCCS UNC NCICU NC Commerce

Partners respond within 14 days.

ALL PARTNERS APPROVE

System drafts for researcher:

- » MOU for relevant data between researcher and relevant partners.
- » Confidentiality agreement.

Researcher returns forms & is notified of final approval.

ALL PARTNERS DENY

Researcher notified of denial.

SOME PARTNERS DENY

System's administrator reaches out to partners to determine reason for denial and seeks unanimity; creates and sends a response to the researcher, potentially suggesting amendments.

SYSTEM EXTRACTS DATA

Automatic Extraction

NCDPI UNC

Manual Extraction

NCCCS NCICU NC Commerce

DATA PROVIDED TO PARTNERS FOR FINAL REVIEW

NCDPI NCCCS UNC NCICU NC Commerce

RESEARCHER NOTIFIED OF DATA AVAILABILITY

Must download data within 10 days.

RESEARCHER SENT DATA DESTRUCTION FORM

Upon completion of research.

The Cabinet worked together to formulate a number of questions they hoped the System would be able to answer, with the goal of turning the data into material that could be used to inform decision-making. In 2019, the Cabinet made a request of NCSchoolWorks to track outcomes of graduates of public high schools.¹⁴ The System successfully provided data from NCDPI, the UNC System, and the North Carolina Department of Commerce. The System was unable to retrieve data from the NCCCS; these data were accessed via the CFS.

Once the data were received from each partner, students were matched by UID. A social security number crosswalk was required to match data from the North Carolina Department of Commerce and NCCCS as the CFS uses social security numbers as the primary identifier. While the request was completed, NCSchoolWorks was not able to produce the data necessary to complete the request without matching data from the CFS.

North Carolina Education Longitudinal Data System

The ELDS is comprised of NCECIDS, NCSchoolWorks, and the CFS. Requests made to this “system of systems” will be completed by matching centralized data from the CFS to federated data from NCECIDS and NCSchoolWorks. North Carolina General Statute Chapter 116E authorizes an ELDS that is charged with facilitating the exchange of individual-level student and workforce data.¹⁵ The statute states that the System operates independent of NCDPI and GDAC but is administratively located within NCDPI. GDAC is tasked with overseeing the drafting and adoption of rules, providing general oversight and direction to the System, approving an annual budget, establishing an advisory committee on data quality, and developing a plan to establish and operate the System.

The ELDS Working Group of the Governor’s Education Cabinet has developed a number of key principles regarding system governance, including that each partner will maintain authority over their data, be able to request data from other partners in the System, and have the right to deny a data request.¹⁶ Additionally, data shared between partners will be identifiable, but data shared outside of the partners must be de-identified.

GDAC and the Education Cabinet continue to move the work of the longitudinal data system forward. The System contributors are currently reviewing a “Contributor Data Sharing Agreement,” which will be able to incorporate future partners as well. The statute also requires that rules be adopted; these rules have been drafted and will be formally considered by the North Carolina Rules Review Committee.

The Working Group has recommended to the Education Cabinet that an independent consultant be engaged, with the charge of working with GDAC to develop a plan to improve interoperability and modernize the ELDS.¹⁷ The consultant will work with each partner, other key stakeholders, agency leaders, and data users to get feedback on the System. The consultant is expected to produce a report by early 2020, which will describe the current state of the System, a vision, objectives, and requirements of the ELDS, recommendations regarding governance, and a sustainability plan.

Challenges Facing the North Carolina Education Longitudinal Data System

- ▶ **Accurately assigning and matching students to UIDs.** UIDs play a critical role in the success of the SLDS, particularly in federated systems. Partners in the System report varying degrees of success managing this process. If a large percentage of students are not assigned a UID, or if a student is not successfully matched to their UID and is assigned multiple UIDs over the course of their academic and workforce transitions, the data produced by the System will be compromised. In order to accurately assign UIDs, each partner entity must have sufficient resources, including both personnel and technology, to manage matching, assigning, and resolving near matches. These resources were found to vary across partners. Postsecondary institutions complete the UID match and assignment process centrally at their respective system office at set intervals rather than locally at an individual campus in real-time. This can cause delays in matching and make it cumbersome to resolve near matches.
- ▶ **Lack of support among partners.** While all NCSchoolWorks partner entities expressed an underlying belief in the value of the data system, there are differences across partners in perceived value of the System to their institution. In some ways, participation in the SLDS was an unfunded mandate. Additionally, a number of the partners engage in bilateral data sharing, resulting in questions regarding what additional value NCSchoolWorks brings to their work. This variability, paired with a lack of clear ownership over the System and internal willingness to commit to data quality, has impacted the progress of the partnership.
- ▶ **Lack of vision about how the System should be structured.** It is unrealistic to expect the System to provide an answer to every question posed. By establishing sets of questions that policymakers, researchers, and institutions want answers to, it would be possible to better cultivate the data from each institution and potentially create standard reports of aggregate data.
- ▶ **No single responsible party to define data points, ensure data quality control, or support appropriate data use.** Partner entities use the Common Education Data Standards (CEDS), and the user interface allows researchers to select data for requests using the CEDS labels and definitions. However, there is no process to ensure that each partner is interpreting elements of the CEDS data dictionary the same way. Because each partner independently determines which data map to each standard, it is possible that these data are not accurately aligned.

It is equally important that someone who knows the data is able to serve as a resource to potential researchers and ensure accurate collection, matching, and use of the data. It is feasible that a data analyst role could be created at GDAC, or that this role could be provided by a third-party group, to ensure that data are being reported accurately.

- ▶ **Confusion about which entity is responsible for sharing data.** An increase in partnerships between K-12 and community colleges to support students who participate in dual enrollment and early college high school programs can make it difficult to determine which entity is responsible for sharing specific data points. Similarly, if a student transfers mid-year, there is a lack of clarity on which partner would report that student's data for that year.

Future of Longitudinal Data Systems in North Carolina

Great progress has been made in building a solid foundation for longitudinal data systems in North Carolina. The contributing partners in the ELDS Working Group are already having conversations about what the next iteration of the data system might be and are thinking ahead to envision what a modern system would look like. There is hope that the ELDS would have a public-facing data dashboard where standard data sets are presented in a meaningful and accessible way. The Community College System is currently working to enhance its UID assignment and matching processes, which would improve the accessibility, accuracy, and timeliness of data they submit to NCSchoolWorks and the ELDS. Questions remain about long-term governance of both systems, though it is evident that a clear governance structure is necessary to continue driving the work forward.

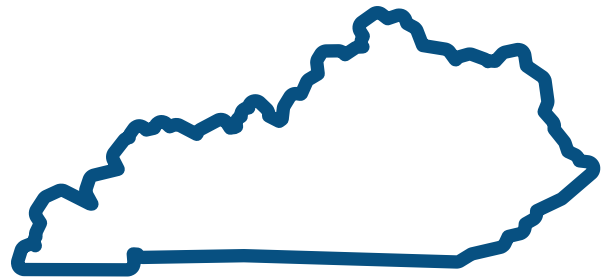
State SLDS Case Studies

While no two statewide longitudinal data systems are identical, there are certainly lessons that can be learned in taking a closer look at how other states have created, legislated, and funded education data systems. Kentucky and Washington both offer examples of SLDSs that successfully provide longitudinal data which informs legislation and policy in their states.

Kentucky

History and Formation

In 2005, the Kentucky Department of Education (KDE) spearheaded the development of a statewide K-12 student longitudinal data system in response to being awarded a \$5.8 million U.S. Department of Education Institute of Education Sciences (IES) SLDS grant. A vendor partner was selected and charged with the task of designing, developing, testing, and implementing the System.¹⁸ The Kentucky Longitudinal Data System (KLDS) was first implemented in mid-2007 as a pilot program and included a limited number of districts.



The System was awarded a second SLDS grant from IES in March 2009 in order to expand the K-12 data available in the System. This grant also allowed for the creation of a “P20 Shared Repository,” which linked K-12, postsecondary, and teacher certification data.¹⁹ Subsequently, the Kentucky P20 Data Collaborative was formed, which included representatives from the KDE, the Education Professional Standards Board, and the Council on Postsecondary Education.

The fallout from the national economic recession in 2008 resulted in the Kentucky state legislature terminating state funding to the SLDS. The KLDS itself lost \$2.2 million in annual funding; additionally, the Kentucky Student Information System lost \$7 million in annual funding. This loss of funding resulted in a reorganization of the System in 2010. In this transition, the P20 Collaborative and its original intent to serve as a single data warehouse changed to a “focus on linking the capabilities of multiple systems.”²⁰

Governor Steve Beshear (2007–2015) was a strong supporter of the state’s ongoing data work and signed an executive order creating the Kentucky Center for Education and Workforce Statistics (KCEWS) in June 2012 to house the KLDS. The KCEWS was ratified into law during the 2013 legislative session. This Center is now called Kentucky Center for Statistics (KYSTATS) and is charged with managing the centralized KLDS, collecting data for the System, and completing data analysis and reports. The System includes data from K-12, postsecondary, and the workforce.

System Overview

The KLDS is a centralized data system managed by KYSTATS, an independent office within the Education and Workforce Development Cabinet. Of the nearly 40 staff members—including analysts, researchers, and developers—nearly half of the staff are working on the KLDS.

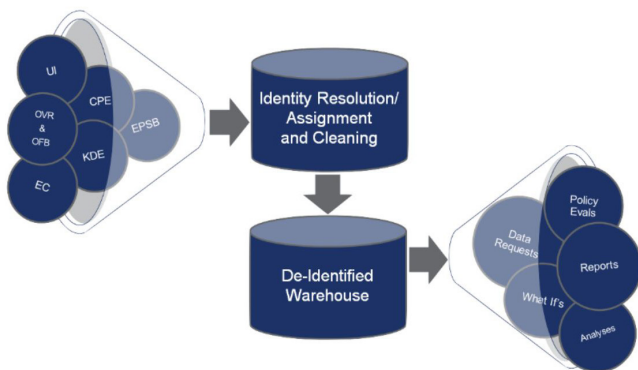
De-identified data are sent to a warehouse from a number of sources including early childhood programs, the K-12 system, postsecondary institutions, and workforce partners. The System relies on an algorithm to link and match the records using social security numbers and other personal identifiers to match K-12 data to postsecondary and workforce data. Personally identifiable information is replaced with unique system identifiers before moving into the de-identified warehouse. Only the Development Team at KYSTATS is engaged in this matching process to ensure data privacy. There are currently 10 years of data in the System, with a delay of approximately six months in real-time data availability. Data are made available

to policymakers, superintendents, researchers, and the public through aggregate reports and in response to specific requests. KYSTATS also provides online dashboards on common topics, such as local workforce areas and academic performance proficiency.

Governance

The center is governed by a board that consists of representatives from the Education and Workforce Cabinet, Council on Postsecondary Education, KDE, and the Kentucky Higher Education Assistance Authority. Effective July 1, 2019, the Cabinet for Health and Family Services will join the board. Each board member recommends an individual from their agency to serve on the Research Committee, which is charged with setting a biannual research agenda for the System. This has served as an avenue for each agency to maintain a sense of ownership of their data as they provide direction about how the data will be used.

Figure 9 | Overview of the Kentucky Longitudinal Data System²¹

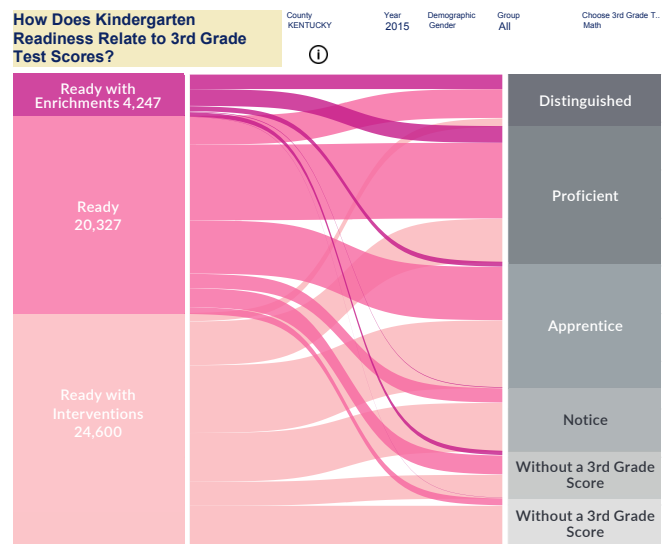


Use and Impact

There are several instances in which longitudinal data made possible through the KLDS have resulted in direct policy impact through increased funding to relevant agencies and programs.

- The statewide initiative, AdvanceKentucky, assesses the impact of AP enrollment and success on student outcomes.²² Results gathered using KYSTATS found that students who completed AP exams and earned a qualifying score were more likely to receive higher college success metrics compared to similar students who did not take an AP exam.²³

Figure 10 | Example of Graphic Available Via KYSTATS Data Dashboard²⁴



- SummerWorks is a program that partners with local businesses, nonprofits, government agencies, and local community organizations to connect youth aged 16–21 with summer employment.²⁵ Data from KYSTATS were used to evaluate the correlation between working during the summer and postsecondary/workforce outcomes for students aged 16–18 and recent stop outs (i.e., students who attended some higher education but then stopped attending before completing a degree) aged 19–21.²⁶ The study found increased employment retention relative to other like-students and some higher results in postsecondary enrollment for stop outs.
- The Kentucky Career and Technical Education Feedback Report is compiled by KYSTATS and used by the Office of Career and Technical Education to create career pathways that align to regional workforce needs.²⁷ KYSTATS compiled occupational projections for the next five years for key sectors defined by the Kentucky Workforce Innovation Board. Using these projections, the Office of Career and Technical Education created the pathways that align to these jobs. KYSTATS then produced a report to show the alignment so that CTE can help schools incentivize high-demand pathways.

Next Steps

KYSTATS continues to work to add additional data to the System. This includes driver's license information to support data matching, corrections, Medicaid, Supplemental Nutrition Assistance Program (SNAP), TANF, job projection data, and a number of workforce data sets including nursing licensure, apprenticeships, and industry certifications.

Keys to Success

A number of factors have contributed to the success of the Kentucky Longitudinal Data System.

- ▶ **Use of a centralized data system** ensures that stakeholders consistently have access to high-quality data that can be pulled in a short amount of time. The centralized system makes it possible for internal and external researchers to continue to build on data sets rather than having to request and complete a new search as research questions evolve. When compared to a federated system, the centralized system ensures that data quality is consistent, as data are cleaned once upon submission to the System, rather than having to be cleaned with each submission following a request.
- ▶ There is **strong partner engagement and support**.
 - › Buy-in across partner agencies has been made possible by the **identification of champions of the SLDS in each agency**.
 - › The System **partners are mandated in legislation**, making expectations for participation clear.
 - › **A strong data governance system** ensures that partner agencies maintain a sense of ownership of their data. This sense of ownership is further promoted by the **required input and sign-off from each partner before a report is released** to the public.
 - › The **creation of tailored reports to meet the individual needs of an agency**, which address specific questions that can only be answered with a longitudinal data system, has allowed partners to see direct benefits from participating in the System.
- ▶ By **publicly publishing a biannual research agenda**, KYSTATS promotes trust and transparency in the use of the data, while also showing what is possible as a result of the data being shared.
- ▶ By **initially piloting the data system in a limited number of districts**, the KDE was able to closely collaborate with offices within the agency to troubleshoot any challenges regarding data sources and matching. Additionally, the limited number of data requests in this early period allowed for the standardization of the data itself, as well as collection and storage processes.

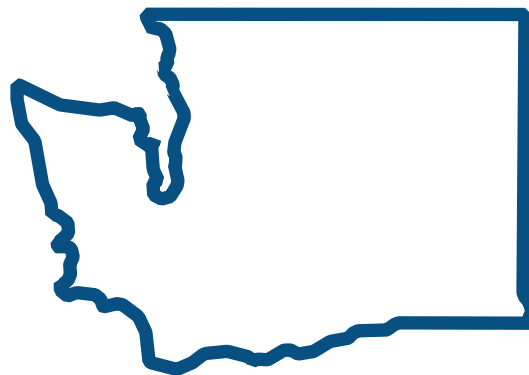
Challenges

- ▶ While KYSTATS receives approximately 30 data requests per month, they continue to be **underutilized by policymakers in the legislature**. KYSTATS is currently developing a plan to engage directly with legislators, demonstrating the types of reports and data available through the System in order to promote legislative understanding and use of the System.
- ▶ The KLDS has long been **reliant on federal grants** to support the System and KYSTATS. Since the first SLDS grants were distributed by the U.S. Department of Education in 2006, Kentucky has received nearly \$19 million in funding. Currently, 88 percent of the center's \$4.3 million budget is funded by federal grants. The remaining portion of their budget comes from state funding, as well as some cost recovery from the completion of external requests.

Washington

History and Formation

In 2005, Washington Governor Christine Gregoire (2005–2013) convened the *Washington Learns* blue ribbon commission to review the state’s education system and make recommendations for improvements across the education continuum. One of the recommendations of the commission was to create a P-20 Council that would, among other things, work to develop a longitudinal student data system.²⁸ The Council was created by Executive Order in 2007 and was charged with tracking student outcomes as they transitioned across the education continuum.



In 2007, the Washington Legislature also passed legislation creating the Education Research and Data Center (ERDC).²⁹ The Center was placed under the Office of Financial Management (OFM) and partnered with the Legislative Evaluation and Accountability Program Committee to conduct analyses of programs across P-20 and the workforce. The legislation required all state agencies involved in education, as well as public postsecondary institutions, to create data-sharing agreements with the ERDC.

In 2009, ERDC was awarded a U.S. Department of Education IES grant to lead the state’s P-20W governance, build a data warehouse, and produce analysis based on critical research and policy questions. This grant also allowed ERDC to expand the System to include early childhood data.

System Overview

The ERDC has a staff of 10 analysts and researchers, as well as additional technical and IT support staff. These positions are funded through a combination of recurring state funds and grant funds from SLDS and Workforce Data Quality Initiative (WDQI) grants. The Center partners with 11 agencies and organizations to offer data from early childhood through the workforce. There are currently 14 years of high-quality data available in the System.

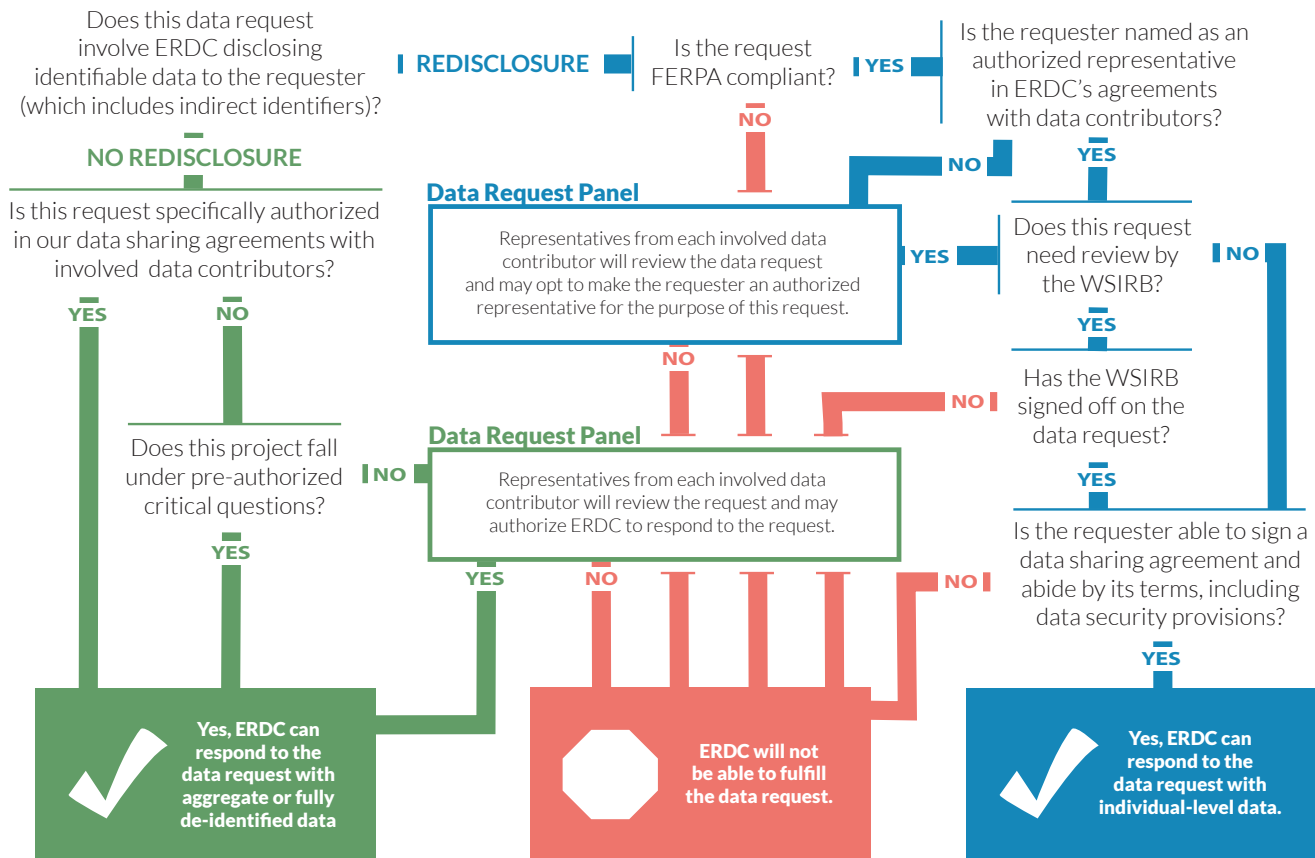
Data are matched using personal identifiers including name, date of birth, gender, and race. The System offers separate processes to complete data requests depending on whether or not the request requires re-disclosure of personally identifiable information. In an effort to build improved transparency and data ownership, ERDC is currently amending its data request process such that partner agencies will be able to opt in to providing data to a given request, whereas previously they had to opt out if they did not want their data to be used. The turnaround time for completing requests ranges from hours to weeks, depending on the complexity of the request and whether pre-existing data sets can be used. The data request process as it currently exists is outlined in *Figure 11* on the next page.

Governance

ERDC differs from other statewide longitudinal systems in that it is not housed within a larger education entity. Instead, it is managed by the Office of Financial Management. The work of the ERDC is guided by three committees that consist of representatives from state-level agencies who share data through the System, stakeholder groups, and ERDC staff members. Each committee serves a different role in guiding the use and functionality of the System. The legislation that guides the work of the ERDC specifies a number of ways the partner agencies must collaborate, including identifying research questions, developing long-term plans to support transition across the continuum, and aiding in the development of an enrollment plan for higher education to meet workforce needs.³⁰

Figure 11 | ERDC Data Request Process³¹

Will I receive data from ERDC?



Use and Impact

Through close partnerships with nonpartisan legislative staff, as well as staff at the agencies, the ERDC works to anticipate the needs of legislators prior to the start of the legislative session, and prepares data sets accordingly. Additionally, a large number of research questions and corresponding data sets have been created, which allows for quick completion of requests. One staff member estimated that the Center receives up to six requests per week during legislative sessions.

Next Steps

The ERDC continues to work to increase the data sources available to the System and is currently working to include vocational rehabilitation training program data, TANF data, and criminal justice data from the Department of Corrections. The legislature recently instructed the Center to include juvenile justice data, which led to the creation of an operational data store at OFM's Criminal Justice Statistical Analysis Center focused on criminal and juvenile justice data broadly. ERDC is also working to identify ways to include social services and healthcare data to help answer research questions about barriers to academic completion.

Keys to Success

A number of factors have contributed to the success of Washington's Education Research and Data Center.

- ▶ **Strong support from the Governor and legislature** has resulted in effective legislation and funding to enable the System to be built and continue meeting the needs of the state. Ongoing legislative funding also instills confidence in the Center's long-term sustainability.
- ▶ By publishing up-to-date information about current requests to the System on their website and taking active steps to make the request process more transparent, the ERDC demonstrates a **strong commitment to transparency**. This helps build trust among the partners and the public. Additionally, the creation of publicly available data dashboards enables the public to see the type of work completed by the Center.
- ▶ After initially receiving only occasional requests, the Center **actively worked to build and maintain a close partnership with the legislature**. This partnership has resulted in meetings with nonpartisan staff to anticipate legislative needs in the upcoming session and allowed the Center to create data sets prior to the beginning of the legislative session to ensure information could be provided in a timely manner once the session began.
- ▶ The ERDC is **housed within a non-education entity**. The ERDC does not manage any education programs and does not regulate any education institutions or agencies. While this was challenging initially, it has likely helped lead to strong support and collaboration with the various partner agencies as the agencies view the ERDC as objective and unbiased.
- ▶ **Use of a centralized data system** ensures that stakeholders consistently have access to high-quality data that can be retrieved in a reasonable amount of time.
- ▶ The **willingness of the Center to continue revising and improving based on best practices** ensures that the System remains efficient, effective, and relevant, and that current data privacy and security standards are met.

Challenges

- ▶ The ERDC continues to work to improve **data matching to support accurate and timely matches across data sources**. The Center is in the process of adding information from state driver's licenses which will help match data more accurately. Matching has also gotten better over the lifetime of the System, as there are now more data points from which to draw matching identifying information.
- ▶ While the legislature has expanded funding for the ERDC and the ERDC has received almost \$27 million in grants from the U.S. Department of Education and the U.S. Department of Labor, **funding has not expanded as quickly as the demands on the Center** have. Additional funding would help the Center meet increasing demand for its services and data products.

Guiding Questions for Policymakers

As policymakers take a closer look at statewide longitudinal data systems, there are a number of questions they should consider when determining the effectiveness and functionality of the Systems. These questions may be helpful in guiding decision-making at the policy level.

1 What questions do we need to answer in order to make progress toward our state's education goals?

Many of the questions that can help inform and shape policy decisions can only be answered by combining data from multiple sources and over longer periods of time.

Determining which questions can be answered by an SLDS and which cannot will require engagement among stakeholder groups, partners, and data analysts. These conversations are best held during the initial planning period of a system in order to frame and set expectations and inform the type of system that will best serve the needs of partners and stakeholders. Determining a research agenda in advance will also guide policymakers as they seek to identify what types of data will be needed, who possesses those data, and which partners should be included as the SLDS is built out.

2 Which audiences will benefit from data provided through the System?

A system should serve policymakers, state-level agencies, system partners, researchers, and families.

Identifying who will be served by the data helps to inform what data are gathered, how the data can best be presented for each group, and what level of staffing support will be required in order to produce and present those data. Each of these audiences can benefit from a strong SLDS for different reasons. Parents can make informed decisions about their child's education, researchers can conduct large-scale analyses to determine causal relationships between programs and their participants, and legislative analysts can determine which policies are producing the greatest benefits for students.

3 What partners are engaged in data sharing?

Policymakers should consider whether all educational entities are sharing data with the SLDS, including early intervention, early childcare, pre-K, K-12, community colleges, private and public four-year postsecondary institutions, and state agencies who own workforce data.

Individuals engaged in building state longitudinal data systems cite that mandating partner engagement through legislation is important in bringing entities to the table. However, legislation alone isn't sufficient to build trusting relationships among partners. Partners should have clear expectations of what qualifies as timely submission of data and processes for data matching to ensure that the System is able to produce complete data sets that are meaningful. Clearly articulated memoranda of understanding (MOUs) play an important role in developing parity among partners. Partners need to be able to see the value in their participation and understand the benefits that their agencies will see from participation.

4 Does the SLDS have a robust governance structure in place to provide necessary oversight and coordination?

The System needs a clear governing body, codified in legislation, which has worked closely with partners in order to build and operate a system that produces timely and accurate data, while ensuring that student privacy is protected.

System governance is critical, and it requires trust and transparency among and between partners. The governing body serves a pivotal role in ensuring that the System is functional. Policymakers should identify whether a clear leadership structure has been established within the governance body to ensure that decisions are made efficiently. This structure can take many forms, but in North Carolina, a lack of clear governance may have contributed to the lengthy completion timeline. The governance structure will also play an important role in considering the benefits and challenges associated with both centralized and federated systems and determining whether a longitudinal data system should be a centralized system or federated system.

Guiding Questions for Policymakers (cont.)

5 What are the staffing and technological requirements needed to reliably collect, maintain, and submit data? What personnel are needed to ensure effective governance and communication?

In order to function effectively, each partner entity must have sufficient human capital to submit, clean, and analyze data. Additionally, the governing body must have analysts to complete quality control checks, process requests, and coordinate partners.

Teams working with state longitudinal data systems should include data analysts who are experienced in analyzing education and workforce data. Experienced education data analysts will ensure that data are analyzed accurately and appropriately so that the data can be used and presented with the appropriate context. This capacity needs to be built both within each member agency and within the coordinating agency. Building the requisite technological ability is costly but critical. Similarly, human capital will need to be added and strengthened at the institution level in order to reliably collect and report data to, and across, systems.

6 How is the System funded, and is that funding sustainable?

The System must be funded through recurring funds from the legislature. Additionally, each partner agency must have sufficient recurring funds to support infrastructure and human capital needs.

The large startup investments made by the federal government enabled many states, including North Carolina, to build systems that otherwise may not have been created. However, many states have not created sustainable funding streams to maintain and improve these systems. In order to ensure the long-term sustainability and success of state longitudinal data systems, there should be permanent or recurring state funding. These funds should be provided not only for the governing body, but also to ensure that there are sufficient funds at each partner entity to maintain the needed technology infrastructure, as well as human capacity both in IT and data analysis. Beyond funding of the overall system, it is critical that the technology infrastructure of each partner is sufficient in order to provide timely, accurate data to the System. This will require varying investments in each partner entity.

7 How will stakeholders access the information in the System?

The System should produce sets of de-identified data that can be analyzed by researchers, regular aggregate reports to serve policymakers, and a public-facing interface with data presented in a way that is easily understood.

For longitudinal data systems to truly be functional, data usability must be taken into account from the beginning. In addition to sharing data with outside researchers, some states have chosen to build this capacity at a central coordinating agency that regularly builds reports and/or interactive dashboards that can be easily used and understood by policymakers and the general public.

8 What steps are being taken to ensure that data privacy is protected?

Efforts to ensure data privacy must be at the forefront of all statewide data efforts.

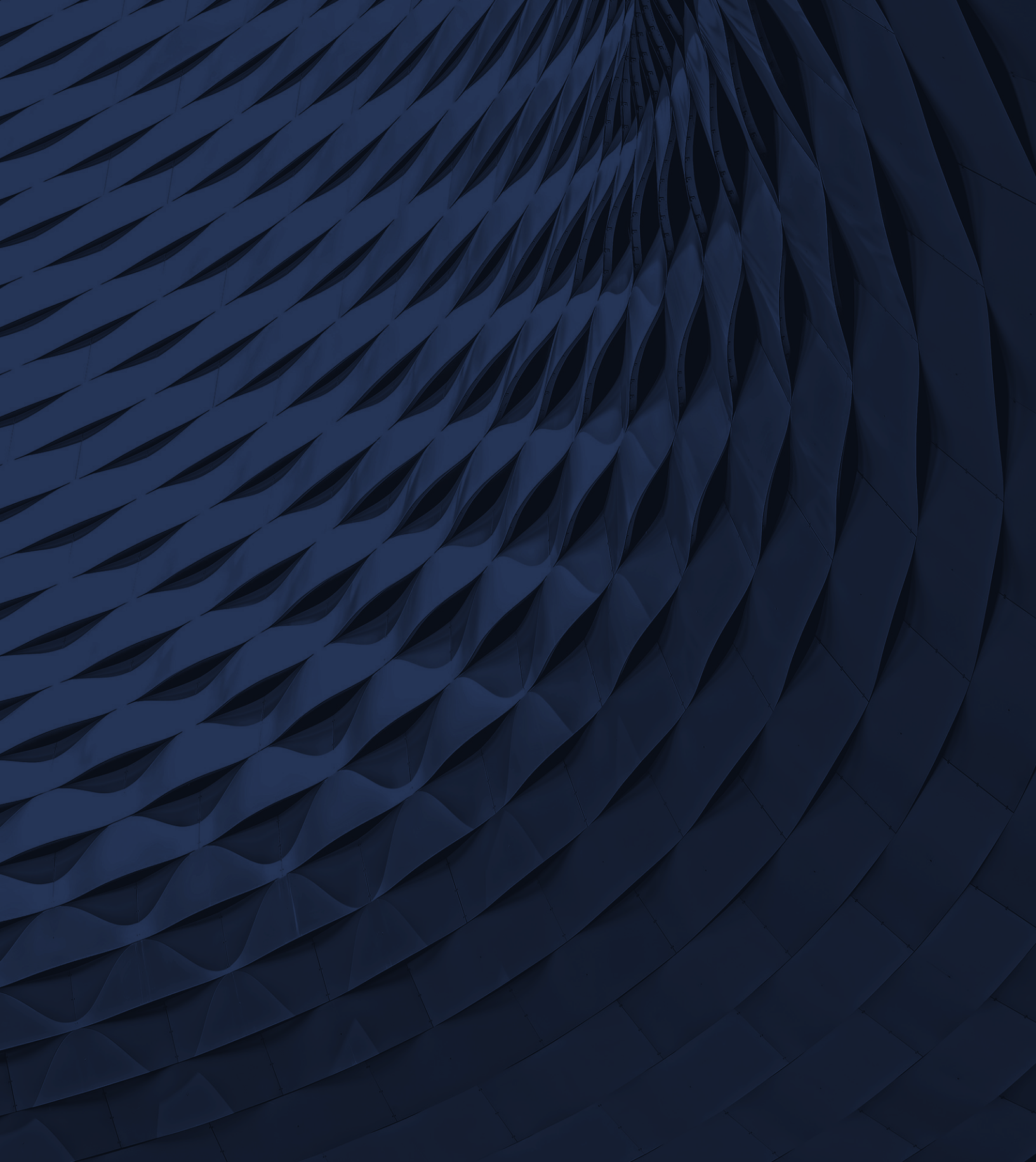
Public trust and confidence in the System is dependent on the knowledge that personally identifiable data cannot be accessed or used to connect an individual to their education and workforce history. Policies must be enacted that span physical, technological, and legal data protections. The governing body plays a critical role in ensuring data are securely held and privacy efforts are transparent to partners, stakeholders, and the community.

Glossary

504	Section 504 of the Rehabilitation Act of 1973	KDE	Kentucky Department of Education
AP	Advanced Placement	KLDS	Kentucky Longitudinal Data System
ARRA	American Recovery and Reinvestment Act of 2009	KYSTATS	Kentucky Center for Statistics
CEDARS	North Carolina's Common Education Data Analysis and Reporting System	LEA	Local Education Agency
CEDS	Common Education Data Standards	mCLASS	A screening tool that measures development of reading skills
CFS	North Carolina's Common Follow-up System	MOU	Memorandum of Understanding
CTE	Career and Technical Education	NCCCS	North Carolina Community College System
ECE	Early Childhood Education	NCDHHS	North Carolina Department of Health and Human Services
ECIDS	Early Childhood Integrated Data System	NCDPI	North Carolina Department of Public Instruction
ELDS	North Carolina Education Longitudinal Data System	NCECIDS	North Carolina Early Childhood Integrated Data System
EOC	North Carolina's End of Course Test	NCICU	North Carolina Independent Colleges and Universities
EOG	North Carolina's End of Grade Test	OFM	Washington's Office of Financial Management
ERDC	Washington's Education Research and Data Center	P-20	pre-K through Higher Education
FERPA	Family Educational Rights and Privacy Act of 1974	P-20W	pre-K through Higher Education and the Workforce
FRL	Free/Reduced Lunch Status	SES	Socioeconomic Status
GDAC	North Carolina's Government Data Analytics Center	SLDS	State Longitudinal Data System
IB	International Baccalaureate Program	SNAP	Supplemental Nutrition Assistance Program
IEP	Individualized Education Program	TANF	Temporary Assistance for Needy Families
IES	U.S. Department of Education Institute of Education Sciences	UERS	North Carolina's Uniform Education Reporting System
IRB	Institutional Review Board	UID	Unique Identifier
IT	Information Technology	UNC	University of North Carolina System
KCEWS	Kentucky Center for Education and Workforce Statistics	WDQI	Workforce Data Quality Initiative

Citations

- ¹ Educational Technology and Assistance Act of 2002, H.R. 3801. 107th Cong, 2nd sess. <https://www2.ed.gov/policy/rschstat/leg/PL107-279.pdf>
- ² “Grantee States.” Statewide Longitudinal Data Systems Grant Program. National Center for Education Statistics. <https://nces.ed.gov/programs/slds/stateinfo.asp>
- ³ Ibid.
- ⁴ “50-State Comparison: Statewide Longitudinal Data Systems.” Education Commission of the States. <https://www.ecs.org/state-longitudinal-data-systems/>
- ⁵ “State Longitudinal Data Systems - All State Profiles. October 2016.” Education Commission of the States. <http://ecs.force.com/mbdata/mbprofallrtan?Rep=SLDS16STA>
- ⁶ U.S. Department of Education. SLDS Topical Webinar Summary Linking K12 Student Data with Postsecondary Data. Washington, D.C.: Institute of Education Sciences, National Center for Education Statistics, 2014. https://nces.ed.gov/programs/slds/pdf/Linking_K12_Student_Data_to_Postsecondary_Data_May2014.pdf
- ⁷ America COMPETES Act of 2007. Public Law 110-69. <https://www.congress.gov/110/plaws/publ69/PLAW-110publ69.pdf>
- ⁸ North Carolina Department of Commerce. Common Follow-up: Information Management System Continuation Review Final Report. Raleigh, NC: Labor and Economic Analysis Division. 2014 https://www.ncleg.net/fiscalresearch/continuation_reviews/FY_2013-14_CR_Documents/CFS-Continuation-Review-2-28-2014.pdf
- ⁹ “About the North Carolina Common Follow-up System.” NC TOWER. <https://www.nctower.com/aboutcfs/>
- ¹⁰ “North Carolina Education Research Data Center.” Duke Center for Child and Family Policy. <https://childandfamilypolicy.duke.edu/research/nc-education-data-center/>
- ¹¹ “Race to the Top - Early Learning Challenge Final Performance Report: North Carolina.” 2016 <https://www2.ed.gov/programs/racetothetop-earlylearningchallenge/2016apr/nc2016finalapr.pdf>
- ¹² Education Longitudinal Data System. Session Law 2012-133. House Bill 964 <https://www.ncleg.gov/Sessions/2011/Bills/House/PDF/H964v5.pdf>
- ¹³ North Carolina General Statute 116E. Education Longitudinal Data System. https://www.ncleg.gov/EnactedLegislation/Statutes/PDF/ByChapter/Chapter_116E.pdf
- ¹⁴ “Longitudinal Data Pilot Research Project.” Education Cabinet Data Workgroup. Prepared by the Labor and Economic Analysis Division of the North Carolina Department of Commerce.
- ¹⁵ North Carolina General Statute 116E. Education Longitudinal Data System. https://www.ncleg.gov/EnactedLegislation/Statutes/PDF/ByChapter/Chapter_116E.pdf
- ¹⁶ “Strengthening North Carolina’s Education Longitudinal Data System (ELDS). Recommendations from the Education Cabinet ELDS Working Group. [DRAFT].” 4/15/2019.
- ¹⁷ Ibid.
- ¹⁸ Kentucky Department of Education. “Kentucky Statewide Longitudinal Data System (KSLDS) Transition Report.” 2011. <https://education.ky.gov/districts/tech/kmp/Documents/KLDSFinalProjectTransitionReportFINAL.pdf>
- ¹⁹ Ibid.
- ²⁰ Ibid.
- ²¹ “KYSTATS Security and Privacy.” Kentucky Center for Statistics. <https://kystats.ky.gov/About/Security>
- ²² Advance Kentucky. <https://www.advancekentucky.com/>
- ²³ Advance Kentucky. “2017 Year in Review: AP College Readiness Program.” 2017. https://www.advancekentucky.com/sites/default/files/inline-files/AdvanceKY_YIR-2017_Dec5.pdf
- ²⁴ “Early Childhood Profile 2019. How Does Kindergarten Readiness Relate to 3rd Grade Test Scores?” Kentucky Center for Statistics. https://kystats.ky.gov/Reports/Tableau/ECP_2019
- ²⁵ SummerWorks. <https://www.summerworks.org/>
- ²⁶ “SummerWorks 2015 Evaluation Executive Summary.” 2015. http://www.clasp.org/sites/default/files/kcews_summerworks2017.pdf
- ²⁷ “The Kentucky Career and Technical Education Feedback Report.” Kentucky Center for Statistics. <https://kcews.ky.gov/Latest/CTE>
- ²⁸ “Washington Learns. World-Class, Learner-Focused, Seamless Education.” November 2006. <http://www.k12.wa.us/Compensation/pubdocs/WALearnsFinalReport.pdf>
- ²⁹ “ERDC’s History.” Education Research and Data Center. <https://erdc.wa.gov/about-us/history-erdc>
- ³⁰ Education Data Center. Revised Code of Washington Chapter 43.41.400. <https://apps.leg.wa.gov/RCW/default.aspx?cite=43.41.400>
- ³¹ “Data Request Process.” Education Research and Data Center. <https://erdc.wa.gov/data-resources/data-request-process>



Copyright © 2019 The Hunt Institute. All rights reserved.